LCG-France Tier-1 & AF

Réunion de Coordination

Fabio Hernandez
fabio@in2p3.fr

Lyon, 15 janvier 2009

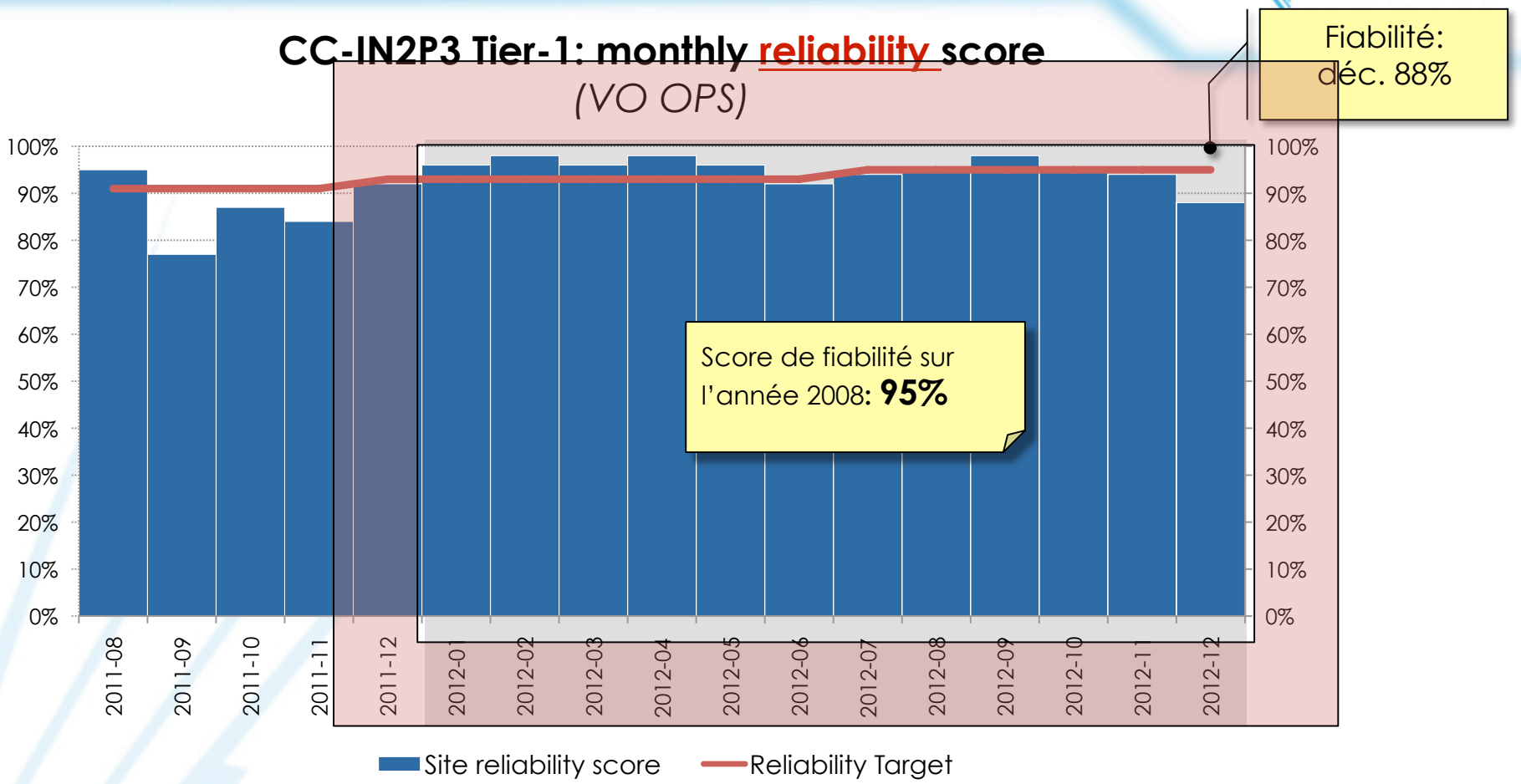
Table des Matières

- Réunions GDB et MB de décembre 2008 et janvier 2009
- Disponibilité, fiabilité, efficacité des sites
- Chantiers en cours
- Thème(s) du jour

GDBs & MBs

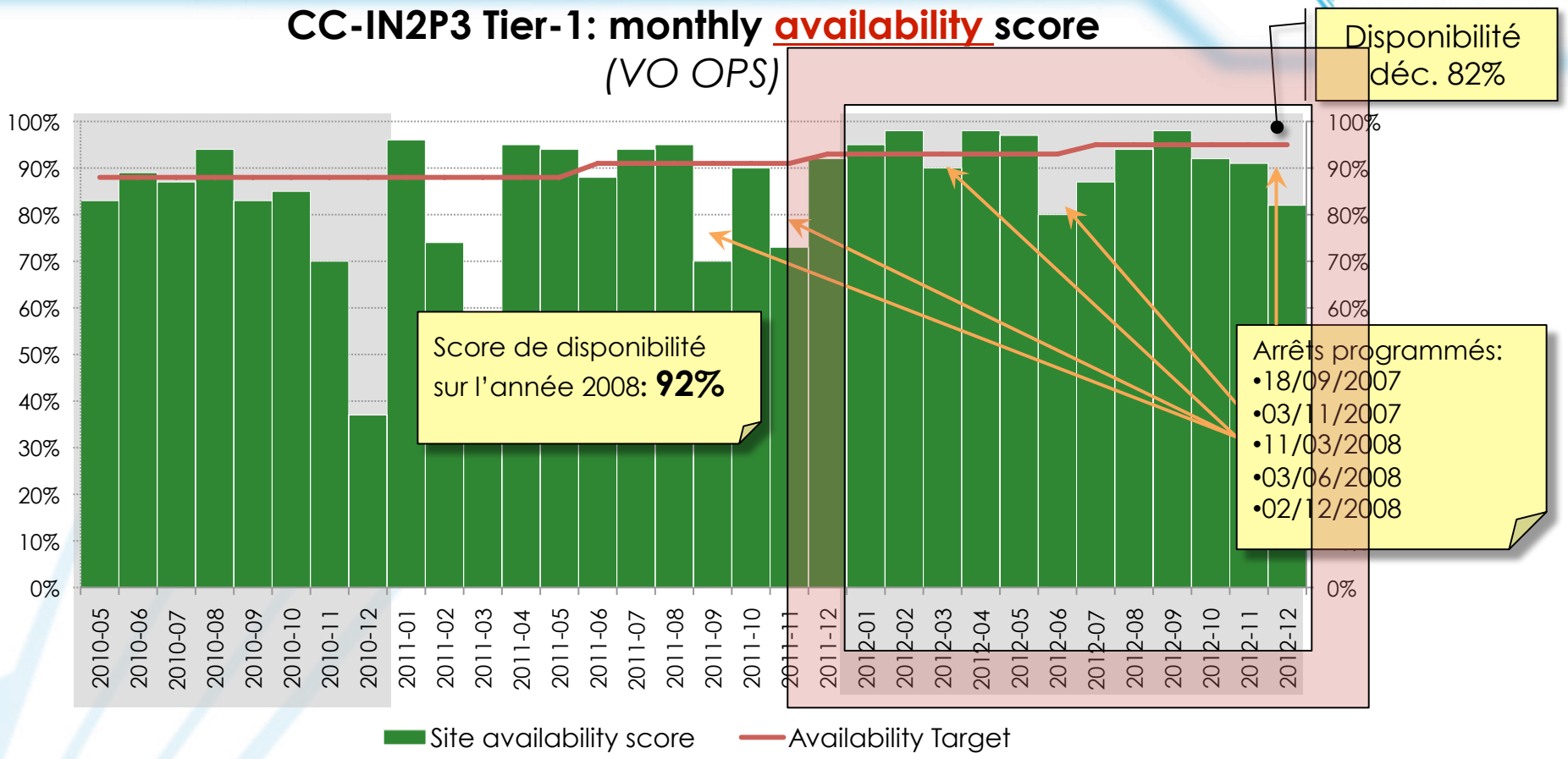
- Agendas:
  - GDB: http://indico.cern.ch/categoryDisplay.py?categId=3l181
  - MB: http://indico.cern.ch/categoryDisplay.py?categId=666
- Principaux sujets traités
  - Mises à jour middleware
  - Fonctionnement des tier-2s
  - Activité récente des expériences

Fiabilité du tier-1



Source: [WLCG T0 & T1 Site Reliability Reports]

Disponibilité du tier-1

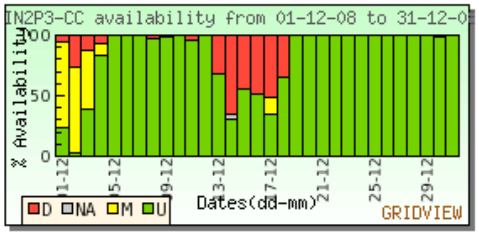


Source: WLCG T0 & T1 Site Reliability Reports

Disponibilité des sites: tier-1 (1/3)

- Décembre 2008

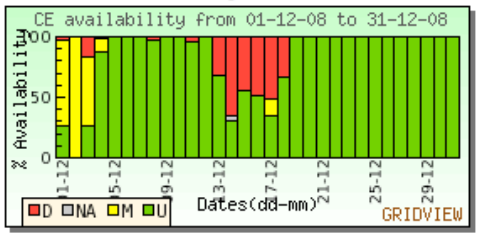
Overall Service Availability for Site:IN2P3-CC VO:OPS (Daily Report)



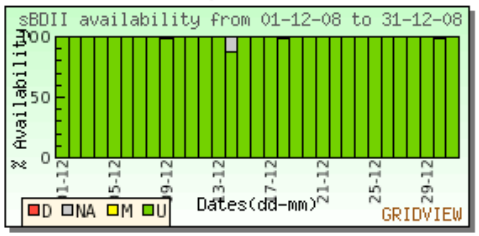
Score: 88%

Individual Service Availability for site:IN2P3-CC VO:OPS (Daily Report)

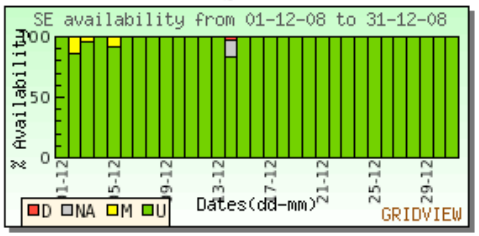
CE



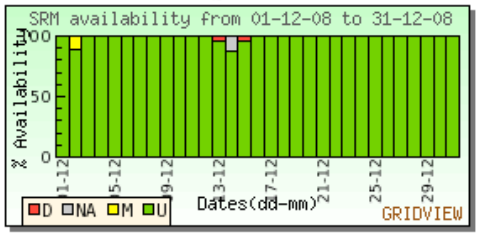
sBDII



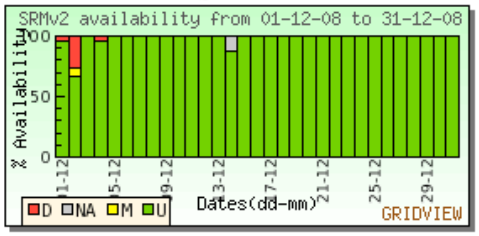
SE



SRM



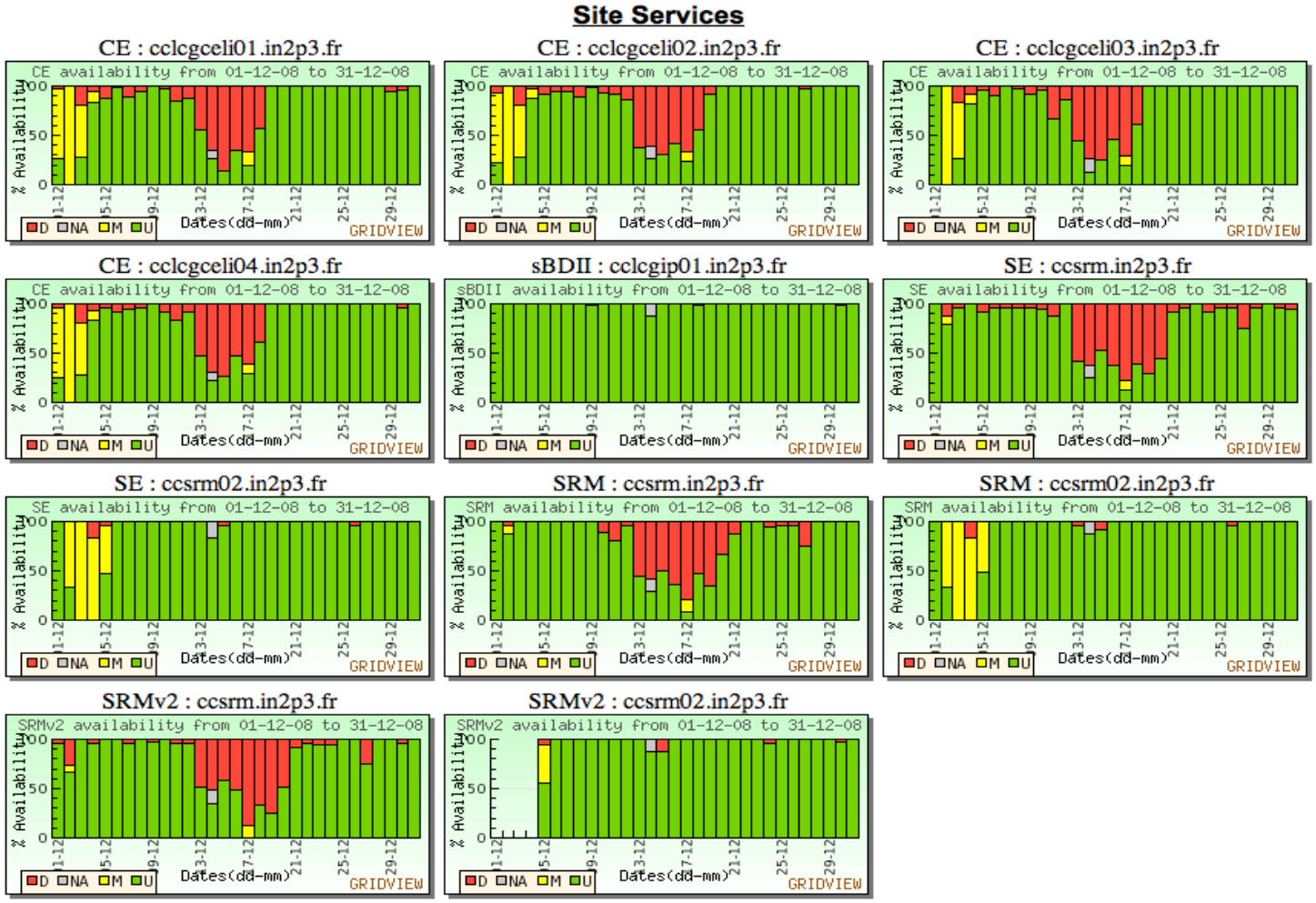
SRMv2



Source: http://gridview.cern.ch

Disponibilité des sites: tier-1 (2/3)

- Décembre 2008 (suite)



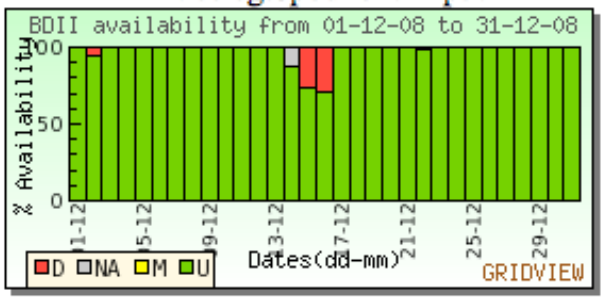
Source: http://gridview.cern.ch

Disponibilité des sites: tier-1 (3/3)

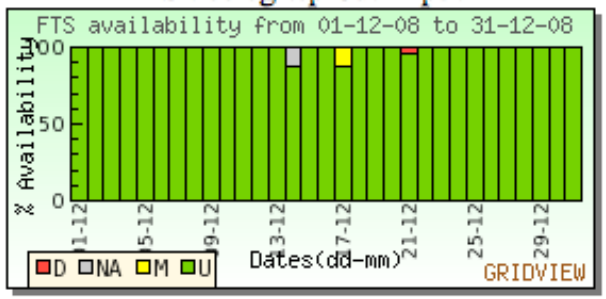
- Décembre 2008 (suite)

Central Services

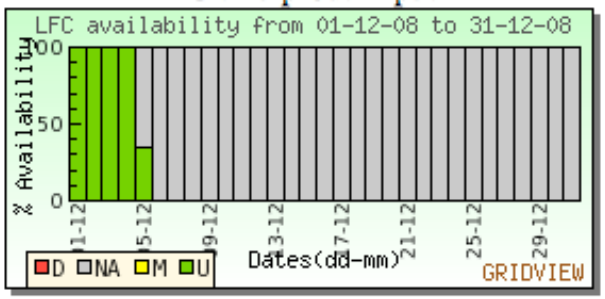
BDII : cclcgtopbdii01.in2p3.fr



FTS : cclcgftsprod.in2p3.fr



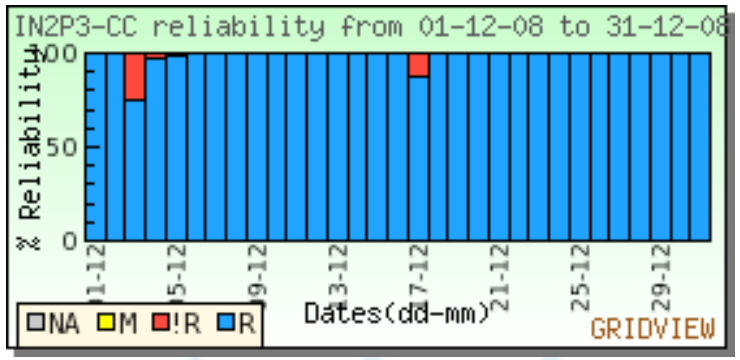
LFC : lfc-prod.in2p3.fr



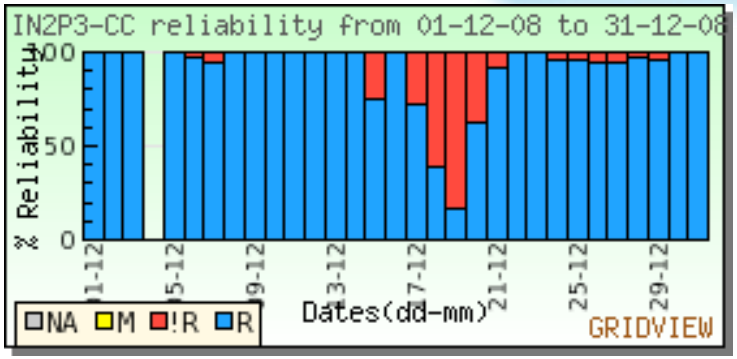
Source: http://gridview.cern.ch

Fiabilité vue par les VOs: Déc. 2008

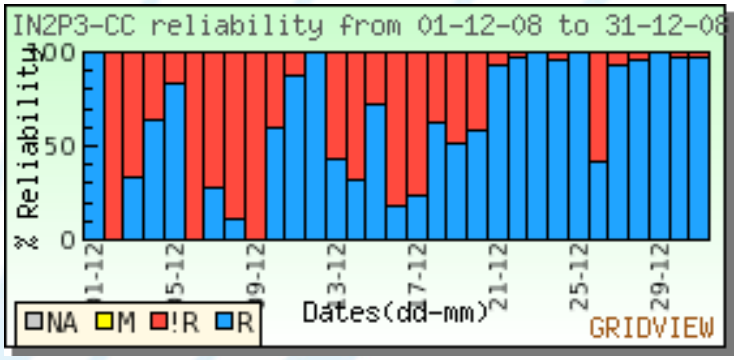
ALICE



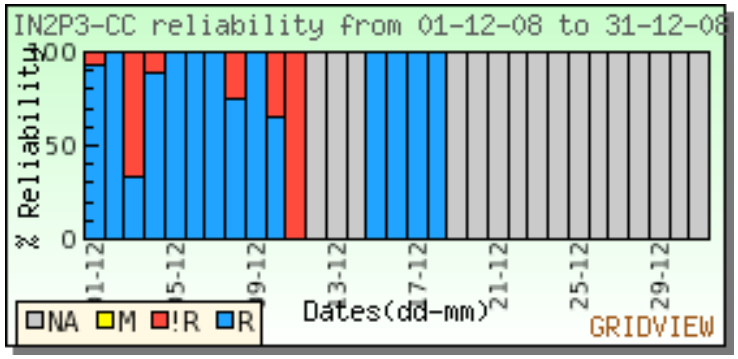
ATLAS



CMS



LHCb



Source: http://gridview.cern.ch

EGEE France – disponibilité & fiabilité

- Décembre 2008

Region	Avail-ability	Reli-ability
UKI	96 %	98 %
CERN	95 %	96 %
SouthEasternEurope	94 %	95 %
CentralEurope	93 %	96 %
NorthernEurope	93 %	94 %
Russia	92 %	93 %
AsiaPacific	92 %	93 %
France	88 %	93 %
GermanySwitzerland	86 %	87 %
SouthWesternEurope	83 %	92 %
Italy	83 %	88 %

Source: https://edms.cern.ch/document/963325

Déploiement des ressources

Background

- Meeting before Christmas with new CERN Director for Physics & Computing
  - Realistic estimates for first collisions for physics ~September 2009
    - Assuming present schedule for re-installation, cool down, testing etc is OK
  - Additional information at that meeting – unlikely that the accelerator would run past mid-November
  - Based on that information it seemed unreasonable and irresponsible to keep to the presently agreed resource deployment schedule
- Subsequently: additional information from CERN management (and mentioned in DG speech to staff and users)
  - Chamonix workshop (Feb 2-6) will discuss 2009 and 2010 running schedules in detail, based on expected capabilities and experiment desires
  - It could happen that there is no winter shutdown this year (but a longer shutdown later) – EDF contract to be retendered; important to do as much physics as early as possible

Ian.Bird@cern.ch 2

What next?

- For 2009 we should relax the requirement to have all installed by April – this is clearly not essential now
  - Could open the possibility to get next generation equipment for the same cost
  - Must be ready in good time for September 2009 (i.e. July/August)
- What else really depends on what comes from the Chamonix discussion:
  - If shutdown in mid-Nov and restart in April we must question the resources that are needed for 2009 and 2010
  - If there is likely to be no (or a much reduced) shutdown then we probably need to keep to the original planning (but with an adjusted or staged deployment)
- Hopefully also to be clarified in Chamonix:
  - Energy, Heavy-ion running
  - Experiments desire to take data (even if they cannot yet analyse it ...)

Source: Ian Bird, Management Board 13/01/2009

Unité de puissance CPU

Transition to a new CPU unit

- Proposal: Take a simple approach and adopt 4,00 as conversion factor.
  - Give more importance to having a simple rule for the transition period than discussing about decimals within the 5% precision.
  - Caveat: during the next days, ALICE and CMS will complete their benchmarking runs up to lxbench08 to confirm the good correlation (no surprises expected).
- Transition period: before end April (Spring C-RRB meeting)
  - Experiments will re-compute requirements tables given the new LHC schedule: new numbers should be computed already with the new unit.
  - Sites should buy SPECcpu2006 and calibrate their farms to report their current CPU power in the new unit. Pledges for the Spring CRRB should be expressed in the new unit.

Nouvelle unité basée sur SPECall_cpp2006

Source: Gonzalo Merino, GDB 14/01/2009

Publication de fermes hétérogènes

Plan For Cluster Publisher

- **YAIM to publish non-overlapping SubClusters**
  - Creation of a new node type - "The Cluster-Publisher"
  - Split the lcg-CE (or cream-CE) into two node types.
    1. Configuration of gatekeeper and GlueCE publication
    2. Configuration of GlueClusters and GlueSubClusters = glite-CLUSTER.
  - The ClusterPublisher can:
    - § CoLocate with the CEnode on small (one CEnode) sites = status quo.
    - § CoLocate with the siteBDII on medium (> 1 CEnode) sites.
    - § Exist alone for massive sites (>> 1 CEnode) sites.
  - ClusterPublisher is a very trivial almost static service.
    - § It publishes static LDIF via BDII describing a site's compute resources.
  - The dynamic exception is the GlueHostApplicationSoftwareRunTimeEnvironment maintained by VO software managers and used by VOs member finding software.
    - § Consideration for software tag publishing must be made to create ClusterPublisher node.

Steve Traylen , steve.traylen@cern.ch , C

RTEPublish Proposal

- **Use apache and mod_gridsite.**
  - Used elsewhere within gLite, e.g WMS.
- **Let apache own all VO tag directories and files i.e**
  - /opt/glite/var/info/<SubClusterUniqueId>/<VO>/*.info
- **We protect each directory via a .gacl file.**
  - A .gacl file can be configured to allow write access for a particular VOMS role.
- **lcg-tags/lcg-ManageTags extended to support either:**
  - Native gridsite commands htcp and htdel.
  - The standard curl client can be used.
- **A working instance has been configured.**

Steve Traylen , steve.traylen@cern.ch , CERN 14

Source: S. Traylen, GDB 14/01/2009

Multiplicité des sources d'information

Problem Statement

- How to reliably find a list of site -> service -> VO mappings in the EGEE production infrastructure

- Current method:
  - Do a 'JOIN' between GOCDB (or SAM) and Top-l contents
  - Given a site:
1. Find list of hosts in the GOCDB for the site
2. Gather from GOCDB the list of services on
3. Lookup GlueService table in BDII for these hosts/services to find supported VOs
   a) And the endpoint specific information (e.g. SAPat Token, Service URI)

EGEE-II INFSO-RI-031688

What to do ?

- Clearly, having a single definitive source of information would solve inconsistencies

Two possibilities:

1. Providing the canonical list of VO mappings and endpoint information in GOCDB

2. Providing the canonical list of services which should be at a site in the Site BDII

EGEE-II INFSO-RI-031688 6

Source: J. Casey, GDB 14/01/2009

Glue v2.0

Proposed Roll Out Plan

Enabling Grids for E-sciencE

1. **Create a hybrid schema file with both v1.3 and v2.0**
   - Deploy across the infrastructure
     - Should have negligible side effects
   - Est. 3 - 6 months after specification fixed
2. **Update information providers**
   - Publish Glue 2.0 information in addition to Glue 1.3
   - Deploy across the infrastructure
   - Est. 4 - 12 months after specification fixed
3. **Update software and tooling as necessary**
   - Est. 6 - 36 months after specification fixed
4. **Remove Glue 1.3 providers when no longer required**
   1. Est. 36 - ?? months after specification fixed

EGEE-II INFSO-RI-031688 16

Source: L. Field, GDB 14/01/2009

Procédures de déploiement du middleware

Post mortem of 2008 releases

- "There is a tendency to bundle too many changes in a release" (CMS)
  - More, smaller releases?
    - Much less efficient for the release process.
    - Would the sites follow?
  - Less change in general?
    - Which changes do we drop?

Actions to take

- Implement well defined rollback procedure (SA3)
- Implement procedure for fast track releases (SA3 & SA1)
- Implement a managed rollout of updates in production (SA1)
- Improve and maintain quality of the release pages (SA3)
- Maximize representation of experiment use cases in certification (SA3)

Rollback

- Current status:
  - One shared repository for all node types
- Planned:
  - "current" repository for every node type (no longer shared) + repositories of the previous update
  - In case of rollback "current" can point to previous update. This can be done per node. Prevents sites of picking up the bad update.
    - Rollbacks are per node type, not per individual rpms.
  - Sites which have already installed the bad update need to downgrade manually.
    - We can provide a recipe
  - The release team needs to sort out how to achieve this from a technical point of view (has implications for our scripts, AFS space, we need to ensure consistency of the repositories etc.)
  - Timeline: end of February 09

6

Source: A. Unterkircher, GDB 14/01/2009

CREAM

Current status of some criteria

- **Condor-G submission**
  - Condor have updated Condor-G to support job submission to the CREAM CE. CMS are now testing this. Plans, progress and results are here: http://hepuser.ucsd.edu/twiki2/bin/view/HEPProjects/CMS-Cream
    - 4-7% failure rate out of 10K short jobs using the "modified" Condor-G
    - Phase 1 (basic functionality testing) completed
    - Phase 2 (further testing on updated CREAM CE) planned for Feb
- **ICE-WMS submission**
  - Now in PPS pilot testing
  - CMS involved now, Alice starting next week
  - On-going stress + functionality testing by certification team
    - Submission rate of 45-50 Hz achieved, but "high" failure rate (Savannah bugs 45437, 45636)
- **Proxy renewal**
  - Two patches to fix this (#2669, #2667). Need to push these patches through testing

WLCG Grid Deployment Board, CERN 14 January 2009

CREAM CE in production

- Currently only 3 instances. Need many more to really shake down the service and see how it performs in a real production environment.
- Have asked the EGEE ROC managers to implement at least one CREAM CE per region in the next few weeks.

Source: N. Thackray, GDB 14/01/2009

Middleware

SL5 m/w clients

- WN/SL5
  - Pilot service now being run via PPS activity
    - ATLAS – OK except for SELinux issue (SL*C* specific?)
    - CMS – OK
    - ALICE – OK
    - LHCB – gssklog issue (not SL5 specific)
  - Pilot service scheduled to end 30th Jan
  - On this basis, we will release the SL5 WN (32/64) at the start of Feb
- Python2.5 support
  - gfal/lcg_utils available via the Application Area
  - DPM/LFC 1.7.1 will come with 2.5 bindings
  - Currently not scheduled to be released as rpms
    - This can be changed
- gcc4.3.2
  - Currently this is not a requirement on middleware as all gcc4.1 binaries will link fine (no C++ involved)

2

Source: O. Keeble, GDB 14/01/2009

SCAS/LCAS/LCMAPS/GLEXEC

- SCAS server
  - Patch #2684 was rejected in December due to memory leaks which mean it could not run for more than 30mins with 2 clients
  - Patch #2719 has been received (11/01/09) and fixes the memory leaks
- lcas/lcmaps
  - Patch #2635 – plugins (verify-proxy and basic)
    - Certification suspended until we have the glexec update
  - Patch #2703 – SCAS client plugin
- Glexec
  - Patch #2381 – updates for deployment on WN
  - Currently with NIKHEF
  - Depends on #2703 (may be merged)

Tape performance

I asked the sites.

a) do you feel your tape system has been stressed by the LHC experiments yet?

b) do you have a method of dynamically seeing the performance of your service to tape?

What Next?

- Information from all experiments on their testing/verification. Including repeated/updated tests.
- Information from sites on their ability to monitor.
- Share information on GDB list to prevent duplicate work.
- Revisit at February GDB
- Present conclusions to LHCC Mini-Review in February.

Source: J. Gordon, GDB 14/01/2009

Événements à venir

- Réunion interne pour établir un plan d'amélioration de l'interaction entre dCache et HPSS
  - En prenant compte l'organisation des données et les patrons d'accès aux données de chaque expérience
  - 30 janvier, 10h-13h, salle 202
  - Agenda détaillé sera disponible prochainement
- WLCG Collaboration workshop
  - 21-22 mars 2009, Prague
  - http://indico.cern.ch/conferenceDisplay.py?confId=16861
- CHEP 2009
  - 23-27 mars 2009, Prague
  - http://www.particle.cz/conferences/chep2009/
- Hepix Spring 2009
  - 25-29 mai 2009, Umeå (Suède)

Aujourd'hui et à venir

- Aujourd'hui
  - Début des rapports de bilan des opérations
  - Organisation et outils de journalisation des opérations des services
- Prochaine réunion
  - 19 février, 13h30-15h30, salle 202
- Réunions à venir
  - http://indico.in2p3.fr/categoryDisplay.py?categId=102

Questions/Commentaires